The School Improvement Mechanic

When speaking to staff who are new to a data-driven approach, it is important to frame conversations in practical ways. Staff might consider approaching performance data in the same way we would want a mechanic to approach our car after it starts making an odd or abnormal noise. We take our car to a mechanic and tell her about the noise. If she were to take the car, *without asking any additional questions*, and tell us she'll have the problem fixed shortly, we'd be concerned. How would we know that our car will be better off after the service call? One would hope she'd probe a little more and ask a series of questions like:

- What does the noise sound like?
- When exactly does the noise occur?
- Does it happen every time you drive?
- Can you pinpoint the location of the noise?
- Does the noise get louder as you accelerate?
- How long has the car been making this noise?
- Does the noise occur in conjunction with braking?
- Does anyone else hear the noise besides you?

We need to put ourselves in the position of the data mechanic. We can gain immense insight into our functions if we ask layered questions like the ones just listed.

Similarly, if a parent received a report stating that his child was not performing up to par, we would hope he would be presented with a detailed explanation of any key terms, the short- and long-term impact of the child's performance, and a synopsis of what the school or district plan is to raise the performance.

Digging a few levels deeper into data than we've traditionally done in the past can be a very revealing exercise. Unfolding as many layers of school performance as possible and making connections among these layers paints a clear map and enables us to chart a course more successfully.

Root cause is typically defined as the most fundamental and underlying reason for a problem or issue. In terms of educational data, it means getting beyond a surface level symptom and exploring down as far as one can go to get to something concretely addressable. If a teacher is told, for example, "You need to really work on your boys' math computation abilities," she hasn't been given much to go on. That directive is so overwhelming and global that it's difficult to formulate a plan of attack. However, if the teacher had been a part of multiple conversations about data where detailed consideration was given, we might reach something more addressable, such as, "We've found that the male students who failed the 'understanding sentence structure' portion of the reading test were the same students who scored poorly on the problem-solving portions of the math test." That thought gives us much more to go on and paints a clear picture of how one might go about dealing with this issue. It also gives us more information to help us figure out who can best help us solve this problem. Is it the principal, the math coordinator, a reading specialist?

Making Connections

Looking at data more deeply to recognize patterns of performance and trouble spots is an ongoing exercise. Being able to explore connections between various sources of data can help pinpoint root causes. For example, can we find out if the subgroup of students who score the lowest in the *Making Inferences* section of a standardized reading test might also have other similar test performance

issues? Did these students all take part in the same remediation assistance the prior year? What about the students who score highest in *Math Reasoning*, for example? Do they take part in school enrichment activities of any kind?

Although it's true that not every angle can be investigated, key student performance issues certainly warrant in-depth consideration beyond one layer. Further, sometimes the most important discoveries cannot be made by looking at a single data set. Although a printout of a class aggregate of standardized test scores gives us great insight into how the class performed as a whole on portions of the test, it doesn't provide much guidance for making instructional changes. We must get into the habit of looking at ongoing data and considering nontraditional sources of information. We always think of test scores, demographics, absenteeism, and so on. But what else might have an impact on how our students are learning, or how successful our teaching efforts are? A 3D team might present the following series of activities as a way to encourage staff to consider all the various data that impacts the learning environment—especially data that extends beyond traditional tests.

Take a moment to read over the data presented in Data Conversation Activity 1 (below). A cursory glance would tell you that the students performed better overall in *Informative Reading* than they did in *Extended Understanding*. Ask any layperson and they would say the same thing. You might also infer that overall the *Higher-Order Questioning* section has the fewest students performing above average. But is that really what this data tells us? Does data in this form help us at all at this point? If we're familiar with the testing items and what exactly they measure, perhaps it does. If we know the curriculum, how this test aligns to that curriculum, and how our own instruction mirrors this curriculum, we might have something to go on. For example, a district-level reading director might immediately begin to think of what standards in the curriculum match these test item categories. In that case, this data might be very helpful in thinking about performance across the district.

However, looking at a generic set of data without more refined headers or other information such as the actual number of items, or the weight given to each category, gives us only a very broad first look at what's going on. It's an important look, no doubt. But because there isn't much in this broad view that gives us something concretely addressable, it's critical that we not stop there.

Data Conversation Activity 1

Take a look at the following sample of district-level subset data. These scores represent district-level aggregate scores from a standardized reading test for sixth graders.

READING COMPREHENSION	% BELOW AVERAGE	% AVERAGE	% ABOVE AVERAGE
Initial Understanding	26	65	9
Vocabulary Reconstruction	22	58	20
Extended Understanding	45	45	10
Informative Reading	21	61	18
Higher-Order Questioning	30	64	6

Let's add a little more information and dig a little deeper into this data set. In Data Conversation Activity 2 (below), we've separated males from females and shown the number of items tested. With raw, clean data that we are sure is accurate, we can make a quick table to examine things a little more closely. We also need to get into the habit of making sure the data we're viewing can provide answers to the questions we have.

Data Conversation Activity 2

MALES

READING COMPREHENSION	*# OF ITEMS TESTED*	% BELOW AVERAGE	% AVERAGE	% ABOVE AVERAGE
Initial Understanding	*8*	20	69	11
Vocabulary Reconstruction	*6*	25	53	22
Extended Understanding	*4*	46	42	12
Informative Reading	*1*	22	60	18
Higher-Order Questioning	*6*	24	66	10

FEMALES

READING COMPREHENSION	*# OF ITEMS TESTED*	% BELOW AVERAGE	% AVERAGE	% ABOVE AVERAGE
Initial Understanding	*8*	24	53	23
Vocabulary Reconstruction	*6*	17	57	26
Extended Understanding	*4*	36	42	22
Informative Reading	*1*	19	61	20
Higher-Order Questioning	*6*	32	56	12

A 3D team member leading a group through these activities might consider the following general questions with regard to the data as it is presented in Data Conversation Activity 2:

- Does this data appear to be accurate? How would we know?
- Would it matter if we knew how many students were tested?
- Can you tell if the number of males and females is equal? Does that matter?
- Is this data in a format that gives us enough information?
- Is this data in a format that is easily understandable?
- What further data do we need, if anything?
- What kinds of assumptions can we make with just the data presented here?
- Does every person in this school or department understand the data as presented?
- Do we all have the same assumptions?

Revisit the data after discussing the questions just listed, and think about these additional issues:

- Is one gender outperforming another overall?
- What specific differences do we see between male and female performance?
- If this chart presented the number of students tested, would that change our assumptions about the data presented here? Why or why not?
- What about differences in performance between subtests?
- To what extent is one group outperforming another?
- Is this difference significant enough to warrant further investigation?
- If we need to dig into this data more, where do we get additional, more detailed reports?
- Can we identify specific students? Is there a need to do this?
- Might we look at this set of data side by side with the two previous years to see if there was growth? Decline?
- Do the scores here paint a complete and accurate picture of what students know? Of what teachers taught?

Even with the added information in the second set of data, we still may not be at the point to make a truly informed conclusion. Considering external factors, testing conditions, and other forms of assessment data might give us a clearer picture. Mixing in previous summative test performance, formative assessment data, and so on might allow us to feel more confident in making a judgment about this data set.

The School Improvement Mechanic Model Revisited

Taking time to ask additional levels of *why*, *what*, *how*, *where*, and *when* can get us closer to discovering the true culprit behind data issues. For example, ask an overworked teacher why her students aren't performing up to par in math, and she may say, "Because kids don't pay attention like they did years ago." Ask her why that is, and she may say, "Because they are so multisensory that they can't just stay focused on the book like I tell them to." Are we satisfied that we've discovered the problem yet? No, there's more to uncover. What else might we want to ask, and how far do we need to dig until we get at least a little closer to what's happening? How can we possibly help this teacher if we stop at her first two responses? Do we want to know what her definition of "focused on the book" means? What if we observed her style of teaching? What if we examined which specific students she teaches? What if we were able to look at some patterns of formative assessment performance in this classroom? What if we could compare this year's scores to last year's and the ones from the year before? How do the students in her classroom perform in other content areas? Does our standard teacher observation system provide enough detail to paint a true picture of why students are performing the way they are? Of the resources made available to this teacher, which does she engage? Why those?

Taking a problem, asking deep, probing questions, and superimposing multiple sources of hard data enables us to determine the root cause(s) of the issue(s) at hand. Being able to identify a problem *at its core* can help us create and enact a tangible strategy to improve on the challenge, making us more certain that our strategy isn't just a Band-Aid for an issue that will continue to fester outside our scope of operation.

Most importantly, moving staff to a level at which each person is constantly in this data discovery, reflection, and action mode ensures that we continuously improve on existing successful efforts, streamline the hard work already being done, and exemplify what we strive to build in our students—lifelong learning.

Victoria Bernhardt in her book *Using Data to Improve Student Learning in Middle Schools* (2004, Eye on Education) notes that "Your best defense against others drawing incorrect or incomplete assumptions about your school is to provide a complete analysis" (p. 238).

Beginning to scrutinize data in new ways, and with new lenses, offers us the chance to finally heal previously hidden fractures that may have plagued our performance for years. Getting into the habit of asking additional questions *each and every time we look at data*, be it small ongoing formative measures, or district-level aggregate standardized test scores, increases the likelihood that we can get to the bottom of academic successes and performance discrepancies.

Excerpted from *Data-Driven Decision Making* by Chris O'Neal (ISTE 2012)